

Yehor Duma

The University of Tokyo
Tokyo, Japan
duma@g.ecc.u-tokyo.ac.jp

Yoshiki Mori

The University of Tokyo
Tokyo, Japan
mori@boz.c.u-tokyo.ac.jp

(In)definiteness-driven Agreement Patterns in Ukrainian Numeral Phrases

Abstract. Slavic numeral phrases (numerically quantified expressions) have been a constant target of inquiry because of tangled agreement relations of a twofold nature. First, the agreement of the cardinal with the complement noun varies for paucal cardinals and higher cardinals – *five* and above. Second, higher numerals exhibit ‘optional’ agreement with the verb, which may appear in singular neuter or plural form. We show that the ‘optional’ agreement varies upon the definiteness and determinacy of the referent of the numeral phrase. In showing that, we present data from the distribution of reciprocal and reflexive pronouns, approximative inversion, and collective nouns with numerals. Our analysis is typologically compatible with research on definiteness and plurality in Western Armenian (Sigler, 1997), plurality and animacy in Turkish (Alexiadou, 2019), and languages marking verbs for definiteness (Ortmann, 2000), contributing to the study of morphological economy and pragmatic effects on agreement (Danon, 2010; Corbett, 2023).

Keywords: numerals; Slavic agreement; semantic agreement; definiteness; plurality.

1. Introduction

Slavic numeral phrases (numerically quantified expressions) have been a constant target of inquiry because of the tangled agreement relations of a twofold nature. First, the agreement of the cardinal with the complement noun varies for paucal cardinals (for example, taking plural nominative in Ukrainian or singular genitive in Russian for cardinal *two* in nominative environment) and higher cardinals – *five* and above. Second, numeral phrases with non-paucal cardinals exhibit ‘optional’ agreement with the verb which can appear in singular neuter or plural form. The latter is attested cross-linguistically: the verb may only ‘optionally’ agree in number with the cardinal-noun construction.

- (1) a. Pjat studentiv bačy{-ly/-lo} veselku.
 five students.pl.gen saw.{past.3p/past.3s.neut} rainbow.acc

‘Five students saw the rainbow.’ [Ukrainian]

- b. xamstaʃshar shanta waʔʕ-et/waʔaʕ-uu.
 fifteen bag-ø fell-F-ø/fall-pl

‘Fifteen bags fell.’ [Lebanese Arabic; Ouwayda, 2017]

- c. vec hyur yega{-v/-n} partiin
 six guest-ø came{-3s/3p} party.dat

‘Six guests came to the party.’ [Western Armenian; adapted from Sigler, 1997, p. 181]

We argue that East Slavic verb morphology agreeing with numeral phrases can be used for definiteness marking in a broad sense in which referential determinacy is also included. We further claim that in cases where the agreement is forced (such as approximative inversion, the presence of reflexives and reciprocals) and not optional, it is due to the syntactic or semantic processing of the definiteness setting. When syntax does not force a certain agreement pattern, the definiteness setting is processed by semantics or discourse. In other words, if the agreement is not required by syntax, contexts in total provide one or another way to resolve verb morphology in languages that permit optional agreement of the numeral phrase with the verb, but a given context does so only one obvious way.

We argue that one type of Ukrainian numerals (which we call collective diminutive numerals, or CDNs) is only fully felicitous in the case their discourse referent, a small animate social group, has been introduced in the preexisting context as a group, and constitutes a case of semantic agreement with the main verb:

- (2) Četvirko studentiv bačy{-ly/ʔ-lo} veselku.
 four.CDN students.pl.gen saw.{past.3p/past.3s.neut} rainbow.acc
 ‘Four students saw the rainbow.’

In Section 2, we start with a discussion of the relationship between agreement and definiteness and come to an intermediate conclusion that definiteness influences numeral-verb agreement across languages. In Section 3, we introduce aspects of Slavic numeral phrase agreement and motivate our choice of Ukrainian as the object of this study. We proceed to claim that the definiteness setting influences the morphology of the main verb: plural marking corresponds to definite numeral phrases, and singular neuter marking corresponds to indefinites. Section 4 provides evidence for this claim in exploring the cases when the numeral-verb agreement is forced, and not optional. We find that the cases of forced agreement plausibly correlate with the (in)definiteness and (in)determinacy of the numeral phrase. Section 5 focuses on collective diminutive numerals in Ukrainian. Here we claim that semantic definiteness effects influence verb morphology when it is not forced by syntax. Section 6 concludes with some prospects. By highlighting definiteness effects present in Slavic numeral morphosyntax and providing novel data, we contribute to the discussion of the interaction between definiteness, plurality, and animacy.

2. Agreement and Definiteness

The puzzle of ‘optional’ agreement with Ukrainian numeral phrases motivates an investigation into the link between agreement and definiteness. We propose that this variation is not optional, but is instead governed by the definiteness interpretation of the numeral phrase in question. This hypothesis is strengthened by data from languages such as Western Armenian and Russian, where a functional correlation between agreement morphology and

definiteness is overtly realized. Precise understanding of this correlation requires a definition of (in)definiteness, which we understand in the sense of Heim (1982) as difference in familiarity between unfamiliar (new) and familiar (known). Definites, then, are connected to their discourse referents which have been introduced earlier. The relationship between definiteness of number phrases and plurality is documented cross-linguistically. For example, Sigler (1997, pp. 167–168) discusses data from Western Armenian. Indefinite subjects in Western Armenian appear with verbs in third-person singular, and definite ones with verbs in plural:

- (3) a. *kəsan* *usanoʁ kənuten-ə-mə* *caxoʁe-c-av*
 twenty student- ɔ exam-abl-indef fail-aor.3sg
 ‘Twenty students failed an exam.’
- b. *kəsan usanoʁ-*(ner)-ə* *kənuten-ə-mə* *caxoʁe-c-an / *-v*
 twenty student-*(pl)-def exam-abl-indef fail-aor-3pl / *-sg
 ‘The twenty students failed an exam.’ [Western Armenian]

Sigler claims that verb morphology is strictly plural when the relevant numeral-noun complex has been introduced into the discourse before the utterance in question takes place. That is to say, the five soldiers in the following sentence have been introduced into the discourse before the sentence is uttered, and the speaker is concerned about their death:

- (4) *ayt baderazm-i-n meč hink zinvor-*(ner)-(ə) əsbann-ve-c-an/*-ab*
 that battle-GEN-dt in 5 soldier-*(pl)-(dt) kill-pass-aor-3p/*-sg
 ‘In that battle the five soldiers were killed.’

If the context is not clear about their definiteness, “soldier” does not require a definite article or plural marker after the numeral. But if the context sets the soldiers as a salient group, both the complement noun and the verb must become overtly plural. The definite article can be sometimes omitted, though.

Pesetsky (1982, pp. 83–84), reviewing earlier traditional grammar works, shows that in Russian, non-agreeing subjects “exclude the group reading” and must be indefinite:

- (5) a. *šest' matematikov* *razlučilis' na mostu*
 six mathematicians.pl.gen parted.3p on bridge
 'Six mathematicians parted company on the bridge.'
- b. *#šest' matematikov* *razlučilos' na mostu*
 six mathematicians.pl.gen parted.3s.neut on bridge
 'Six mathematicians parted company on the bridge.'

However, with enough context, (5b) above becomes felicitous in Russian. For example, there is a fad in mathematical community to part (or end a relationship) on one of the Königsberg bridges to signify that there can be no solution to the relationship. The craze faded but some six of the community did perform the ritual. To simplify somehow, (5b) has a reading that it holds for a specific group of six mathematicians that each small group of them partitioning (or covering) the whole group one way or another can be predicated of with part-on-the-bridge (some sort of cumulative reading). In each small group, there is at least one member of the group of the six mathematicians. In an extreme case, the predication with part-on-the-bridge comes down to each individual of the group of six mathematicians (parting from unmentioned persons; some sort of distributive reading). Wider context, such as knowledge of the world, routinely influences agreement (see Danon, 2010 for examples from Hebrew).

The link between agreement and definiteness, observed in Western Armenian and Russian, provides a framework for analyzing more complex phenomena. Accordingly, the next section turns to 'optional' agreement patterns found with numeral phrases in Slavic. We will begin by surveying the key parametric variations across the Slavic family to contextualize the issue and motivate our focus on Ukrainian. The Ukrainian data will then serve as the basis for our central claim: that this agreement variation is not optional, but is determined by the (in)definiteness and (in)determinacy of the phrase.

3. Slavic Numeral Phrases and 'Optional' Agreement

The morphosyntax of Slavic numeral phrases is notoriously convoluted, with significant variation across Slavic languages. In this section, we briefly

introduce the chief riddles of Slavic numeral morphosyntax and justify our choice of Ukrainian as the object of this study.

What all Slavic languages without articles share is ‘genitive of quantification’. Cardinals *five* and above assign plural genitive to the complement noun:

- | | |
|-------------|---|
| (6) a. Sedm | lékařů |
| seven | doctors.pl.gen [Czech, West Slavic] |
| b. Sedam | liječnika |
| seven | doctors.pl.gen [Croatian, South Slavic] |
| c. Sem' | vračej |
| seven | doctors.pl.gen [Russian, East Slavic] |

However, in East and West Slavic, for higher cardinals (but not for other types of numerals, such as collective or diminutive numerals) which are in inherent or lexical (assigned in connection with θ -marking or lexically selected) case positions (7a), we observe ‘genitive override’ (Bailyn, 2004). It means that genitive plural on the complement noun is ‘overridden’ and receives case from the verb that assigns an indirect case to both words in the numeral phrase. This is a property of constructions with cardinals *two* and above, but not for full-fledged nouny cardinals, such as numeral bases of *thousand* and above. If a Ukrainian verb such as *keruvaty* (‘to govern, to operate’) or *pyšatysa* (‘to be proud of’) that always assigns instrumental case requires an object that is a numeral phrase, both the cardinal and the complement noun receive instrumental, and the genitive of quantification on the complement noun is overridden. Despite the surface similarity, the three traditionally assumed branches of Slavic languages diverge in oblique case positions: we observe ‘genitive override’ in East and West Slavic, but not in BCS. Genitive override happens exclusively with the numeral phrases with cardinals in inherent or lexical case positions as other constructions that feature genitive (like possessive and partitive, see (7b)), and other types of numerals (like (7c)) are never overridden:

- | | | | | |
|-----------------|---------------|---------------------|---------------|----------------------|
| (7) a. Profesor | mehaniky | pyšajetsa | simoma | studentamy |
| professor | mechanics.gen | prides.3s.pres.refl | seven.inst | student.masc.pl.inst |

‘The mechanics professor prides himself in (his) seven students.’

- b. Profesor mehaniky pyšajetsa **avtomobilamy** **studentiv**
 professor mechanics.gen prides.3s.pres.refl car.masc.pl.inst student.
 masc.pl.gen

‘The mechanics professor prides himself in his students’ cars’
 (...that they had reconditioned)

- c. Profesor mehaniky pyšajetsa **dvijkom** **studentiv**
 professor mechanics.gen prides.3s.pres.refl two.CDN.inst student.
 masc.pl.gen

‘The mechanics professor prides himself in (his) two students.’

Finally, crucially relevant to this paper is ‘optional’ agreement, only present in East Slavic languages. It has received ample attention in literature, most notably in Franks (1994). His contribution was to show that previous analyses failed to account for the diversity within Slavic languages. In Polish and BCS only non-agreeing option is accepted within the standard norm:

- (8) a. Pięć czarownic przyjecha{-lo/*-ły}
 five witch.gen.pl arrived{neut.sing/3pl}

‘Five witches arrived.’ [Polish; adapted from Lyskawa (2020, pp. 10–11)]

- b. Devetnaest aviona {je prešlo / *su prešli} granicu
 nineteen plane.gen.pl {is flown-across.sg/areflown-across.3pl} border.acc

‘Nineteen planes crossed the border.’ [BCS; adapted from Bošković (2013, p. 70)]

- c. Pjat studentiv bačy{-ly/-lo} veselku.
 five students.pl.gen saw.{past.3p/past.3s.neut} rainbow.acc

‘Five students saw the rainbow.’ [Ukrainian]

Even within East Slavic, Ukrainian seems to have the most diverse paradigm of patterns of agreement because of several types of numerals derived from nouns: diminutive numerals (DNs: *dvijka*, *trijka*), collective numerals (CNs: *dvoje*, *troje*) and collective diminutive numerals (CDNs: *dvijko*, *trijko*):¹

¹ Emmerick (1992) notes that in forms such as *dvaká-* and *triká-* in Old Indian, the suffix *-k-* functions as diminutive. Nevertheless, Comrie (1992) observes, Modern Bulgarian

Table 1. Agreement properties of Slavic numeral phrases

	Genitive of quantification	Genitive override	Optional agreement
BCS	Yes	No	No ²
West Slavic	Yes	Yes	No
East Slavic	Yes	Yes	Yes

Several analyses have been proposed to account for ‘optional’ agreement to show that agreement is never really optional in these cases. Recall the Lebanese Arabic data from the introduction, repeated below:

- (10) xamstaʃshar shanta waʔʃ-et/waʔaf-uu.
 fifteen bag-ø fell-F-ø/fall-pl
 ‘Fifteen bags fell.’ [Lebanese Arabic; Ouwayda, 2017]

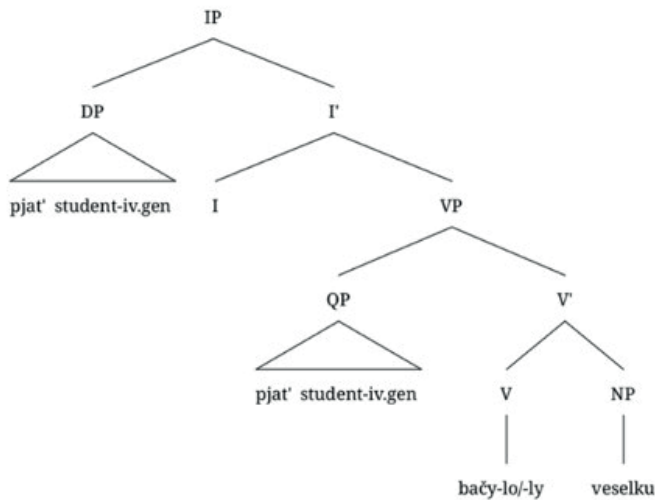
Notice absence of number marking on the complement noun of the numeral phrase. When the complement noun is marked, the form of the verb will always correspond to it, as Ouwayda (2017, p. 267) states: we “[...] can rule out the possibility that plural agreement is simply optional [...], since it is obligatory whenever the noun is overtly plural-marked”.

Ouwayda’s proposal is the “optional pluralizing projection #P” to tie plural agreement with collectivity. The numeral can be in Spec#P to require plural agreement, and in SpecQP for singular agreement (from Ouwayda 2017), with the collective reading available only for cardinals in Spec#P.

A similar proposal has been made by Franks (1994) for explaining optional agreement by movement, with the difference that the whole phrase moves:

- (11) Pjat studentiv bačy{-ly/-lo} veselku.
 five students.pl.gen saw.{past.3p/past.3s.neut} rainbow.acc
 ‘Five students saw the rainbow.’ [Ukrainian]

² We thank an anonymous reviewer for pointing out that the categorical classification here requires nuance. While optional agreement in BCS is largely considered ungrammatical or is highly dispreferred, it is not entirely unacceptable for all speakers. This marginal acceptability is often analyzed as an instance of semantic agreement. For detailed discussion, see Franks (1994) and Bošković (2006, 2024). Another reviewer observes that, in sentences like (8b), optional agreement may occur in colloquial usage when the complement noun is animate; however, elicitation with our informants yielded inconclusive results.



There are problems with such approaches: Ouwayda's account of optional agreement is contingent on the complement noun unmarked for number (described as having a [-pl] feature) as well as a pluralizing projection. Franks' account of optional agreement crucially hinges on different categorial status and height for numeral phrases. Differently from them, Bošković only assumes categorial difference. Bošković (2024) criticizes Franks' approach of different structural heights because the elements in SpecVP can bind subject-oriented anaphors. Instead, Bošković accounts for the optionality by analyzing numeral phrases as quirky subjects: full verbal agreement is triggered when the phrase is a Nominative-cased DP, whereas default (non-agreeing) morphology appears when it is a caseless QP.³

Our main claim is that (in)definiteness setting correlates with the number marking on V: plural for definites (and referential determinacy), and default (3s.neut) for indefinites (and referential indeterminacy). We will present as evidence cases where a specific agreement pattern with the verb is required. Our approach does not contradict Bošković's, as it is known that there is a connection between QP/DP distinction and the (in)definiteness setting (Xiong & Huang, 2020). We will further present evidence that, giv-

³ For an alternative analysis of East Slavic numeral phrase structure – especially in relation to paucal numerals and the resulting case and number mismatches – see Pesetsky (2013), who argues for the presence of a **Quant** projection.

en a context, there is usually one preferred way of resolving the agreement. We infer that when syntactic agreement is underdetermined, semantic factors such as presupposition and collectivity/distributivity distinction on the nominal reference can license verb morphology, in line with the Agreement Marking Principle (Wechsler, 2011).

4. Forced Agreement and (In)definiteness

Slavic languages are known for ‘free’ word order, with the basic word order being SVO (see (12a)): there, the object stays in situ, and the subject is raised to specTP. The next most common word orders are OSV and OVS (as in (12b)), usually derived by Topicalization (A-bar movement), where a fronted object highlights an alternative:

- (12) a. Sim studentiv bačy{-lo/-ly} veselku
 seven students.pl.gen saw{neut.3s/3pl} rainbow.acc
 ‘Seven students saw the rainbow.’
- b. Veselku bačy{-lo/-ly} sim studentiv
 rainbow.acc saw{neut.3s/3pl} seven students.pl.gen
 ‘As for the rainbow, it was seen by seven students.’

Word order variation of this kind does not influence agreement in most Slavic sentences. For example, genitive override persists, and optional agreement is also preserved, even when the word order is scrambled. However, in one case, word order inside the numeral phrase leads to only one grammatical option, namely, in the case of approximative inversion construction (AIC). This is the first case in this section of placing semantic restrictions on word order and agreement, discussed below in 4.1. Other cases of forced agreement will concern the presence of demonstrative, presence of reflexive *self* and reciprocal *each other*, and semantically motivated forced agreement with collective diminutive numerals. We will show that agreement in all of these cases has to do with the (in)definiteness of (in)determinacy of the referent. The aim then is to generalize the effect of definiteness to optional agreement.

4.1. Indefinites are Singular: Approximative Inversion

When the cardinal and the noun appear in reversed order, they are said to have an ‘approximative reading’ (13). The cardinal receives the role of range approximation, or uncertainty operator (Zaroukian, 2010). Nothing else in East Slavic can be inverted with the approximative effect:

- (13) Studentiv sim bačy{-lo/*-ly} veselku.
 students.pl.gen seven saw{neut.sing/3pl} rainbow.acc
 ‘Around seven students saw the rainbow’ [Ukrainian]

If adjacency of the cardinal and the complement noun is violated, approximative interpretation is not possible (14a, b). As a result of internal merge, no head can intervene between the inverted words (14c), an observation made by Billings & Yadroff (1996). More precisely, if a head intervenes between the numeral and the common noun in the base positions (14b), internal merge from inside of this phrase is not possible to build an approximative inversion (14c):

- (14) a. Studentiv ja bačyv sim
 students.pl.gen I saw.past.1s seven
 ‘As for the students, I saw seven (of them)’
 #’I saw around seven students.’ [Ukrainian]
- b. sorok veselyx student-ov
 forty happy.pl.gen students.pl.gen
 ‘Forty happy students’ [Russian]
- c. (*veselyx) student-ov (*veselyx) sorok (*veselyx)

The consistent link between this inversion and an indefinite, approximative reading suggests a syntactic basis for the effect. If we posit that the numeral phrase exists within a larger DP projection, we are able to propose that this construction is headed by a functional element encoding indefiniteness (e.g., a [-definite] feature). The noun-cardinal inversion can then be analyzed as the syntactic mechanism that satisfies, or ‘checks’, this feature. This framework provides a principled explanation for the restriction on the ap-

proximative inversion that it goes together with the verbal agreement of the third person singular neuter (see the above example (13)). As a conjecture, we can predict that, since the structure is syntactically marked as indefinite, no element that requires or imposes definiteness (such as a demonstrative discussed below) can precede the inverted noun-cardinal complex, as this would result in an ungrammatical feature conflict (see the next section).

In certain contexts, the numeral phrase can be accompanied by a classifier-like general noun, sometimes referred to as a pleonastic or measure noun. In Russian, for example, the noun *čelovek* (lit. ‘person’) can be used as a sort of classifier after a number, taking a complement common noun (15b). If the approximative inversion is built up from this construction, it is the classifier-like general noun *čelovek* (15a), not the complement common noun (15c), nor both together (15d), that undergoes inversion with the cardinal.

- (15) a. *Čelovek* *sem’* *studentov* *vide{-lo/*-li}* *radugu*
 people.pl.gen seven students.pl.gen saw{neut.sing/3pl} rainbow.acc
 ‘Around seven students saw the rainbow.’ [Russian]
 [DP [D *čelovek*_i [-def]]][CLP [Num *sem*][CI’ [CI $\emptyset$ _i] [NP *student-ov*]] ...
 (CI = Classifier)

- | | | |
|----------------------|------------------|------------------|
| b. <i>sem’</i> | <i>čelovek</i> | <i>studentov</i> |
| seven | people | students |
| c. <i>*studentov</i> | <i>sem’</i> | <i>čelovek</i> |
| students | seven | people |
| d. <i>*čelovek</i> | <i>studentov</i> | <i>sem’</i> |
| people | students | seven |

By analyzing the classifier-like general noun *čelovek* as the head of the entire nominal expression in (15a), we immediately capture why it controls external verbal agreement, triggering the default singular neuter form. Locating the [-def] feature responsible for the approximative semantics in the same D head formalizes the mechanism of inversion as a feature-checking operation between the D head and its QP complement.

This analysis is aligned with the broader observations on the nature of indefiniteness. Billings & Yadroff connect approximative cardinality with

“some” indefinites, such as “Some men were in the garden”. For us, approximation and indeterminacy that arises from this construction points to the fact that even if a group of the students can be identified, individual students cannot. Indeterminacy about the number easily transfers to anonymity of individuals and indeterminacy of reference because of its fluent demarcation. However, further research is necessary to fully understand AIC.

Noun-numeral inversion is linked to indefiniteness cross-linguistically. It has been shown that in Shupamem (Cameroon), $N > \text{Num}$ order only has definite interpretation, and $\text{Num} > N$ is exclusively indefinite, which is the opposite of approximative inversion construction in East Slavic (Vázquez-Rojas et al., 2011). Similar effects are reported by Dryer (2008) for Nias (Austronesian). If the “default” singular neuter agreement is correlated with indefiniteness, one expects to find it in conflict with contexts requiring definite interpretation, which is discussed in the next section.

4.2. Definites are Plural: Demonstratives, Reciprocals, Reflexives

In this section, we will go one step further and show that pronouns occupying the D position above the numeral uniformly determine agreement with the verb depending on the pronoun’s morphological form. First, demonstratives must be referential and lead to a definite interpretation. We therefore predict that the verb agreement with third person singular neuter is not acceptable. The prediction holds:

- (16) Ci / *ce pjat studentiv vyrišy{*lo/-ly} rozlučytys’
 these/(*this*) five students.pl.gen decided.past.{3s.neut/3pl} to-disband
 ‘These five students decided to disband.’

We see that approximative inversion construction (AIC) and demonstrative pronoun force different agreements on the verb. AIC has an indefinite interpretation, and demonstrative pronouns force a definite interpretation. The pronoun also carries a presupposition of existence and uniqueness. If only one element can occupy the D position, the prediction arises that approximative inversion cannot appear with demonstratives (in fact, nothing can precede AIC, as shown above):

- (17) a. #Ci studentiv sim bačy{-ly/-lo} veselku
 these.pl.nom students.pl.gen seven saw{neut.sing/3pl} rainbow.acc
 #‘These around seven students saw the rainbow.’

b. The (*around) seven students saw the rainbow [English]

The ungrammaticality of these examples stems from a direct feature conflict within the DP. Under our analysis, the head of the DP is specified for a [def] feature. Overt lexical items like the demonstrative *ci* inherently carry and assign a [+definite] value. In contrast, the AIC is the syntactic realization of a [-def] value, licensed by AIC. The two are therefore in competition to set the feature of the same head. The constructions in (17) are ungrammatical because they attempt to value D as both [+def] and [-def] simultaneously, resulting in a failed derivation.

Secondly, reciprocals are words like ‘each other’. They naturally require a plural subject, otherwise the sentence is ungrammatical (Heim et al. 1991). Agreement data from the distribution of Ukrainian numeral phrases seems to support it. In East Slavic higher numeral phrases, the presence of reciprocal ‘each other’ forces a plural agreement to the verb.

- (18) Sim studentiv zna{* -lo/-ly} odyn odnogo
 seven students.pl.gen knew{neut.sing/3pl} one one.gen
 ‘The seven students knew each other.’

If our claim is correct, the subject as the antecedent of reciprocals is interpreted as definite or determined. At this moment, recall the Königsberg-bridge example above (5). There, the singular neuter morphology was possible if, applied to the group reference, the distributive reading down to an intermediate (i-sum) level and further down to individuals was allowed. In the case of thorough distributive reading, some covert reference to outside of the group was considered if the verbal predicate was collective. In the case of reciprocals here, the commonality is the distribution down to an intermediate (i-sum) level, but the distributive reading down to the individuals in the group and the accompanying covert reference to outside of the group are ex-

cluded. Conditions for the referential indeterminacy are still incomplete.⁴ It is predicted, then, that similarly to the case involving a demonstrative pronoun, reciprocals are semantically unlikely to cooccur with approximative inversion (19a), but they do cooccur with plural demonstrative pronouns (19b), because no agreement conflict arises. The predictions seem to hold:

- (19) a. #Studentiv sim zna{-lo/-ly} odyn odnogo
 students.pl.gen seven knew{neut.sing/3pl} one one.gen

‘(Some) seven students knew each other.’

- b. Ci sim studentiv zna{* -lo/-ly} odyn odnogo
 these seven students.pl.gen knew{3s.neut/3pl} one one.gen

‘These seven students knew each other.’

Finally, the same extends to reflexives⁵ as well:

- (20) Sim studentiv bačy{* -lo/-ly} sebe
 seven students.pl.gen saw{neut.sing/3pl} self

‘Seven students saw themselves.’

- (21) a. #Studentiv sim bačy{-lo/-ly} sebe
 students.pl.gen seven saw{neut.sing/3pl} self

‘(Some) seven students saw themselves.’

⁴ If we follow the judgment of Pesetsky on (5), the reasoning is simpler: predications on the group reference allow for the plural morphology, ones down to an intermediate (i-sum) level allows for the singular neuter. We leave it open which line we should follow for further research.

⁵ The relationship between reflexives and reciprocals is not well understood and has a great deal of cross-linguistic variation. Büring (2005, pp. 220–221) observes that “plural reflexive sentences cannot be interpreted as reciprocal; for example, they like themselves cannot mean that they each like the other(s)”. In German and many Slavic languages, however, they allow for reciprocal interpretation:

- (i) a) Sie mögen sich
 they like self
 ‘They like each other.’ / ‘They like themselves.’ (German; Büring, 2005, p. 221)
 b) Oni siebie lubią
 they self like.pl
 ‘They like each other.’ / ‘They like themselves.’ [Polish]

b. Ci	sim	studentiv	bačy{* -lo/ -ly}	sebe
these	seven	students.pl.gen	saw{neut.sing/3pl}	self

‘These seven students saw themselves.’

Our analysis rests on the assumption that the definiteness and determinacy of a numeral phrase can be formally determined within the syntax. The section has provided strong support for this view by examining three distinct contexts where verbal agreement is unambiguously plural. The unifying factor across these cases is that each construction requires a referentially determined antecedent, which we have modeled with a [+def] feature that syntactically forces plural agreement. These instances stand in contrast to cases where the syntax leaves the [def]-value underspecified, resulting in the agreement optionality that is ultimately resolved at the semantic or discourse level. Thus, we have established a crucial correlation: when the definiteness of a numeral phrase is syntactically fixed, so too is its agreement pattern. In the next section, we look at a special case of strongly preferred agreement in Ukrainian after briefly describing general agreement patterns of Ukrainian numeral phrases.

5. Semantic Agreement, Definiteness, and Collective Diminutive Numerals (CDNs)

It is well known in the literature that φ -features of words can be expressed even if syntactic agreement is not possible, as in “donkey anaphora” (Preminger, 2019):

(22) No linguist who has purple pants looks silly in them/*it

In “donkey sentences”, the plural agreement is present despite absence of c-command. Further, the nominal in question (*purple pants*) is in the relative clause, which means that features are not syntactically interpreted. Donkey anaphora type of agreement is then semantic and is based on semantic φ -matching. Slavic nouns referring to humans (such as *profesor*) have an interpretable gender feature. Although the noun is morphologically masculine, it can be semantically feminine. Syntactic agreement fails here, and we observe semantic agreement:

These nouns, to the best of our knowledge, have never received a proper analysis in the literature, including in traditional grammar. The attention to Ukrainian diminutive collectives was mentioned as a separate phenomenon in 1932 in *Narysy z ukrajins'koji syntaksy* (*Essays on Ukrainian Syntax*) (Smerečynskyj, 1932, p. 99). However, the discussion is quite brief, and the author mentions that “there are diminutive numerals like...četvirko, pjatirko etc. not present in other languages”. The most authoritative dictionary of Ukrainian (Bilodid, 1970) does not describe *dvijko* as diminutive, but rather as an ‘immutable group’ form, and gives the following examples:

- | | | | | | |
|---------|--------|-------|----------------|-------------|--------------|
| (26) a. | Kolo | stiny | stojat | dvijko | divčatok |
| | around | wall | stand.pres.3pl | two.CDN.nom | girls.pl.gen |

‘Two girls are standing near the wall.’ (Vasylchenko, 1954).

- | | | | | | |
|-------------|-----|-------------|------------|-------------|------------|
| b. Pered | nym | vidčajdušno | tancuva-ly | dvijko | cyhančat |
| in-front-of | him | boldly | danced.3pl | two.CDN.nom | roma-kids. |
| | | | | | pl.gen. |

'Two Roma kids boldly danced in front of him.' (Hončar, 1954).

No other types of numerals prefer plural agreement on the verb: diminutive numerals such as *dvijka*, *trijka*, etc. are common feminine nouns and agree in all ϕ -features, collective numerals such as *dvoje*, *troje*, etc. exhibit optional agreement, and cardinals *five* and above exhibit optional agreement. Based on the discussion of syntactically forced agreement driven by definiteness and determinacy, other examples of semantic agreement in East Slavic, and properties of CDNs that characterize them as collective we claim that they introduce a semantic definiteness and determinacy effect on agreement with the verb. Unlike the cases of forced agreement in the previous section, this effect does not stem from a formal syntactic definiteness feature. Rather, the semantic mechanism at play is one of group formation and discourse salience similar to the reaction to the collectivity and distributivity for the Königsberg-bridge example and to the antecedent formation of reciprocals and reflexives. By lexicalizing a small number of individuals as a single, bounded ensemble (e.g., ‘a twosome’), the CDN presents its referent as a uniquely identifiable and cohesive entity within the discourse model, thereby motivating the plural agreement typically associated with definite, specific groups.

6. Conclusion and Future Problems

This paper has argued that variable patterns of verbal agreement with numeral phrases in East Slavic are not a matter of optionality, but rather serve as a sensitive diagnostic for the definiteness and referential determinacy of the subject DP. We claimed that the (in)definiteness feature in D correlates with the number marking on V: plural for definiteness and determinacy, and default (3s.neut) for indefiniteness and indeterminacy. Definiteness-driven agreement in this sense is syntactic in cases in Section 4 but semantic with CDNs (and hence can be overridden). Semantic agreement is driven by the context as CDNs cannot be licensed without the discourse referent introduced before. Discussion of definiteness-driven agreement is typologically compatible with research on definiteness and plurality in Western Armenian (Sigler, 1997), plurality and animacy in Turkish (Alexiadou, 2019), and languages marking verbs for definiteness (Ortmann, 2000), such as Hungarian, contributing to the study of morphological economy and pragmatic effects on agreement (Danon, 2010; Corbett, 2023). However, great issues still remain. For example, it is not clear what happens to the definiteness setting when the numeral phrase is in the oblique cases and the genitive on the complement noun is overridden, as the genitive of quantification is the only genitive that can be overridden in East and West Slavic languages. Bošković (2006) claims that the numeral in oblique case is an AP. We end up with a cardinal-noun construction that can be NP/DP, QP or AP depending on the construction in question. Considering semantic ϕ -matching of gender features in Slavic discussed above, plural agreement on verbs in presence of CDNs may also be extraneous, especially because it can, in rare cases, be overridden by a singular demonstrative pronoun. Other remaining issues include the sources of approximative meaning and its connections to indeterminacy, issues in morphology of numerals and roles of suffixes (in particular the apparent polysemy of the suffix *-k-*), and distributivity and collectivity in numeral phrases.

REFERENCES

- Alexiadou, A. (2019). Morphological and semantic markedness revisited: The realization of plurality across languages. *Zeitschrift für Sprachwissenschaft*, 38(1), 123–154.
- Bailyn, J. F. (2004). The case of Q. In *Formal approaches to Slavic linguistics* (Vol. 12, pp. 1–26). Ann Arbor: University of Michigan Press.
- Billings, L., & Yadroff, M. (1996). Phrasal phonology prosodic correspondence in syntax: Russian approximative inversion. In R. Kager & W. Zonneveld (Eds.), *Phrasal phonology* (pp. 45–72). Dordrecht: Foris.
- Bošković, Ž. (2006). Case and agreement with genitive of quantification in Russian. In C. Boeckx (Ed.), *Agreement systems* (pp. 99–120). Amsterdam: John Benjamins.
- Bošković, Ž. (2013). Phases beyond clauses. In L. Schürcks, A. Giannakidou, & Ž. Bošković (Eds.), *The nominal structure in Slavic and beyond* (pp. 116–158).
- Bošković, Ž. (2024). On (partially) quirky subjects, numeral subjects, and subject-oriented anaphor binding: Nominal and non-nominal subjects and their structural positions.
- Büring, D. (2005). *Binding theory*. Cambridge: Cambridge University Press.
- Comrie, B. (1992). Balto-Slavonic. In J. Gvozdanović (Ed.), *Indo-European Numerals*. Berlin: Walter de Gruyter.
- Corbett, G. G. (2006). *Agreement*. Cambridge: Cambridge University Press.
- Corbett, G. G. (2023). The agreement hierarchy and (generalized) semantic agreement. *Glossa: A Journal of General Linguistics*, 8(1), Article 35.
- Danon, G. (2010). The definiteness feature at the syntax-semantics interface. In A. Kibort & G. G. Corbett (Eds.), *Features: Perspectives on a key notion in linguistics* (pp. 144–165). Oxford: Oxford University Press.
- Dryer, M. S. (2008). Order of numeral and noun. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online* (pp. 362–365). <http://wals.info/chapter/89>
- Emmerick, R. (1992). *Old Indian*. In J. Gvozdanović (Ed.), *Indo-European Numerals* (Vol. 57, pp. 163–198). Berlin: Walter de Gruyter.
- Franks, S. (1994). Parametric properties of numeral phrases in Slavic. *Natural Language & Linguistic Theory*, 12(4), 597–674.
- Heim, I. (1982). *The semantics of definite and indefinite noun phrases* (Doctoral dissertation). University of Massachusetts Amherst.
- Heim, I., Lasnik, H., & May, R. (1991). Reciprocity and plurality. *Linguistic Inquiry*, 22(1), 63–101.
- Lyskawa, P. (2020). The structure of Polish numerically-quantified expressions. *Glossa: A Journal of General Linguistics*, 5(1), Article 49.
- Ortmann, A. (2000). Where plural refuses to agree: Feature unification and morphological economy. *Acta Linguistica Hungarica*, 47(1–4), 249–288.
- Ouwayda, S. (2017). On the DP dependence of collective interpretation with numerals. *Natural Language Semantics*, 25(4), 263–314.
- Pesetsky, D. (1982). *Paths and categories* (Doctoral dissertation). MIT.
- Pesetsky, D. (2013). *Russian case morphology and the syntactic categories* (Vol. 66). MIT Press.
- Preminger, O. (2019). *The Anaphor Agreement Effect: further evidence against binding-as-agreement*. Unpublished ms., University of Maryland at College Park.
- Sigler, M. (1997). *Specificity and agreement in standard Western Armenian* (Doctoral dissertation). MIT.

- Vázquez-Rojas, V., Myler, N., & Wood, J. (2011). Definite and indefinite numeral phrases in Shupamem. *New York University Working Papers in Linguistics (NYUWPL)*, 3, 231–244.
- Wechsler, S. (2011). Mixed agreement, the person feature, and the index/concord distinction. *Natural Language & Linguistic Theory*, 29(4), 999–1031.
- Xiong, J., & Huang, C. R. (2020). QP for indefiniteness: With special reference to Sinhala and Chinese. *Lingua*, 236, 102772.
- Zaroukian, E. (2010). Approximative inversion revisited. In J. Bailyn, E. Dunbar, Y. Kronrod, & C. LaTerza (Eds.), *Proceedings of the Annual Workshop on Formal Approaches to Slavic Linguistics* (pp. 146–160). Ann Arbor: Michigan Slavic Publications.
- Smerečynskij, S. 1932. *Narysy z ukrajinskoji syntaksy (u зв'язku z frazeologijeju ta stylistykoju)*. Kharkiv, Rad. škola
- [Смеречинський С. 1932. *Нариси з Української Синтакси (у зв'язку з фразеологією та стилістикою)*. Харків: Рад. школа]